**顺序状态的文字逻辑结构**

1,进行文字顺序逻辑分析的目的就在于拆解文本,但是传统拆解模式过于依赖非标准化的语言结构.我还是希望能够在顺序结构上实现数组排列相关,而非通过语义结构.以该结构为基础考虑按数组陈列特点构成字词序列模式产生相关的文段结构.

2.很简单,首先是拆解,分类整理与用法定义,根据用法重构.最终去实现文章结构的再组合.

有很多变量,具体实现需要以获取的字词的逻辑叠加顺序为准.于是一期拆解完成前.之后的顺序排列模式方案无法完全确定.但是就有限的范例来看,应该可以实现.复杂性上的覆盖目前因为人工取样实验的样本有限.

于是

1. 范围选择
2. 样本选择
3. 一级拆分(单字)(同时进行单字赋值,以用法为基础进行文本内的有限叠加).
4. 文本内有限叠加(然后以文本扩充为基础继续叠加)为基础词之后继续按照词序叠加构成高级词.
5. 当完成相对文本内的高级词叠加之后,最终确认文本内的有效词(例如高能带电粒子为一个词,不能拆解为三个低级词\_高能,带电,粒子),有效词汇确认完毕之后再进行该部分的相对该部分样本文本进行有效词部分的词序顺序分析.由单字扩充到高级词的时候可以确认字词一级别的相对有效性(对于文本内有效词而言)
6. 对于特殊词的问题,这里非指的是特殊名词,而是文字本身的特殊字词,例如的,之,以及关联词.
7. 有效词需要确定该词汇的唯一标准(不一定唯一)构型

例如质子-质子链反应作为中文标准词构型,亦可模糊对位为:

质子质子链反应(符号残缺),

质子链反应(内容残缺),

质子质子链(内容和符号问题),

阳子-阳子链反应(日文汉字直接使用造成的跨语种表示差异问题),

阳子阳子链(跨语种,符号缺失,内容缺失),

阳子链反应(跨语种,符号缺失,内容缺失)

等多种结构的非标准表述,构成标准到非标准的各种映射.

对于多个标准构型对应一个模糊词汇暂时不做讨论,因为该部分的非标准化省略部分很可能夹杂在其他部分当中.例如上下文意义(同样具备相对完善的文字顺序结构模型的话可以获得词汇短语级别的初级复杂关联,然后利用这种相对可靠的复杂关联再进行堆叠.获得整句文段模式的终极复杂关联,以此再堆叠出章节乃至全文的高级复杂关联.

文本内的信息关联,文本内信息方向的学科扩展知识关联.文本内的字形关联(火星文),文本内的发音关联(空耳),

音译结构可以利用国际音标IPA进行语言间的发音相似性分析并确定对音准则.

意译结构在当下条件下很难具备条件(因为汉字的数字结构模式和拼音文字完全不同,日文里面又有假名发音对应的汉字问题s)

创意工坊.修正设计对口商家.(从选择用料,设计等等全覆盖.版权申请交涉)